Measuring the potential for self-connectivity in global air transport markets: implications for airports and airlines

Citation for published version:
Suau-Sanchez, P, Voltes-Dorta, A & Hector, R-D 2016, 'Measuring the potential for self-connectivity in global air transport markets: implications for airports and airlines', *Journal of Transport Geography*, vol. 57, pp. 70-82. https://doi.org/10.1016/j.jtrangeo.2016.09.013, https://doi.org/10.1016/j.jtrangeo.2016.09.013

Digital Object Identifier (DOI):
10.1016/j.jtrangeo.2016.09.013
10.1016/j.jtrangeo.2016.09.013

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal of Transport Geography

MEASURING THE POTENTIAL FOR SELF-CONNECTIVITY IN GLOBAL AIR TRANSPORT MARKETS: IMPLICATIONS FOR AIRPORTS AND AIRLINES

Pere Suau-Sanchez (corresponding author)
Cranfield University
Centre for Air Transport Management
MK43 0TR, Bedfordshire, United Kingdom
p.suausanchez@cranfield.ac.uk

Augusto Voltes-Dorta
University of Edinburgh Business School
Management Science and Business Economics Group
EH8 9JS Edinburgh, United Kingdom
avoltes@becarios.ulpgc.es

Héctor Rodríguez-Déniz
Statistics Division
Department of Computer and Information Science
Linköping University
581 83 Linköping, Sweden
hrodriguez@becarios.ulpgc.es

ABSTRACT

One of the strategies that air travellers employ to save money is self-connectivity, i.e. travelling with a combination of tickets where the airline/s involved do not handle the transfer themselves. Both airports and airlines, particularly low-cost carriers, have recently started catering to the needs of this type of passengers with the introduction of transfer fees or the development of self-connection platforms. The evidence provided by the existing literature, however, suggests that the degree of implementation of these strategies falls short of its true potential. In order to investigate how much self-connectivity could be observed in global air transport markets, this paper develops a forecasting model based on a zero-inflated Poisson regression on MIDT data. We identify the airports that have the highest potential to facilitate self-connections, as well as the factors that hinder or facilitate the necessary airline agreements at major locations. The results from this paper have many implications in regards to the widespread implementation of self-connection services and the future of the air travel industry.

Keywords: Self-connectivity; air transport networks; connection builder; Poisson regression.

---

1. INTRODUCTION

In recent times, low-cost carriers (LCCs) have experienced increased competitive pressures. They have become less able to exploit new sources of latent demand in mature air transport markets and have suffered from increasing route density problems (de Wit and Zuidberg, 2012). LCCs have also been affected by the reconversion of holiday charter operators into scheduled airlines with business models that are relatively close to the low-cost paradigm (Dobruszkes, 2013). In this context, a widely reported process of "hybridization" has occurred, with many LCCs adopting several characteristics of traditional full-service operators (Klophaus et al., 2012), with whom they now compete more directly (Morrell, 2005). These characteristics include, price bundling, codesharing agreements, and the operation of connecting flights on top of their traditional point-to-point services (Morandi et al., 2015; Fageda et al., 2015). In Europe, Vueling and Air Berlin were among the first LCCs to start offering flight transfers in their hub at Barcelona and Palma de Mallorca respectively. Another example is Norwegian, which charges a connecting fee primarily to cover the costs of transferring passenger baggage between its own flights. Globally, Air Arabia (UAE) and

Citilink (Indonesia) are known for offering flight connections at their main bases (ICAO, 2014). The provision of these services shows that LCCs are interested in generating new sources of revenue by catering to the needs of self-connecting passengers, who design their own flight itineraries outside the boundaries of traditional airline connectivity with the objective to save money. These passengers travel on a combination of multiple tickets and take care of their own baggage transfers since the airline/s involved do not handle the transfer themselves. This definition includes both inline LCC transfers as well as other types of interline connections (ViaMilano, 2016). In some LCC airport bases is common to see price-sensitive passengers sleeping in the terminal to self-transfer between late-evening and early-morning flights. Some well-known cases are London-Stansted, Bergamo and Bordeaux airports.

A few European airports have also shown interest in tapping this segment of demand (Fageda et al., 2015). For example, London Gatwick and Milano Malpensa implemented new platforms to facilitate self-connections and improve travel experience. In exchange for a fee paid during the reservation process (available from online booking platform/s that participate in the scheme), self-connecting passengers at Gatwick are offered, at the very least, a baggage transfer service as well as insurance against the risk of missing their onward flight in the event of delays (Gatwick Airport, 2015). Gatwick is indeed a special case, and the success of these programmes depends critically on airlines signing up to these services (in order to increment the potential number of self-connecting itineraries covered at each airport)[1]. From airlines' perspective, the advantage of participating in such schemes may be linked to the potential benefits that interline connectivity brings in terms of economies of traffic density (Starkie, 2007).

From the airport's perspective, a direct benefit is linked to increased non-aeronautical revenues generated by the extra connecting passengers (Malighetti et al., 2008). However, they can also benefit in terms of route development, particularly if self-connections help improving the way in which short-haul low-cost frequencies feed passengers to long-haul flights, thus making international routes more sustainable in the long term (Fageda et al., 2015) and potentially developing the airports' position as international gateway. In relation to this, the CEO of Gatwick Airport argues that their "diverse short- and long-haul airline mix" improves the airport's suitability to host a self-connection platform (The Independent, 2015). Another argument in support of that idea is the development of low-cost long-haul services (De Poret et al., 2015) that effectively expands the scope of LCC self-connectivity to long-haul markets as new opportunities arise for passengers to find cheap fares in these longer routes. The negotiations between Ryanair and Norwegian (who operates long-haul frequencies) in regards to possible collaboration clearly indicate that airlines are keen to develop those travel opportunities (CAPA, 2016).

In spite of the above, the fact that self-connectivity is primarily a passenger strategy to save in airfares indicates that price-sensitive leisure travellers would remain the key target for this type of self-connection platforms (Burghouwt, 2007; Maliguetti et al., 2008; Fageda et al., 2015; OAG, 2016)[2]. This fits nicely with the kind of passengers and destinations that LCCs all over the world typically focus on: short-to-medium-haul holiday markets (Dobruszkes, 2013). Hence, it is not surprising to find many short-to-medium-haul Mediterranean destinations among the self-connecting routes promoted by Gatwick and Milano airports even in markets

[1] The airlines that have signed up to Gatwick Connects service are (September 2016): Aer Lingus, Air Europa, British Airways, easyJet, Flybe, Meridiana, Mornarch, Norwegian, TAP, Thomas Cook, Virgin Atlantic, WestJet and WOW Air.

[2] Nevertheless, the characteristics of the self-connecting passenger may vary depending on the market and region. For example, the promotional video of GatwickConnects seems to target young couples and business travellers. In this regard, more research is needed to fully understand the different profiles of the self-connecting passenger.

that are well served by direct frequencies or traditional flight connections (See e.g. Gatwick Airport, 2015; Via Milano, 2016).

The existence of both short/medium and long-haul dimensions of self-connectivity suggests that a complex assessment of an airport's potential to develop this type of traffic platforms is warranted, yet still not available in the literature. Depending on an airport's geographic location and airline mix, different patterns of specialization in self-connecting routes could appear. In addition, while the existing literature on the topic of self-connectivity is scarce (Malighetti et al., 2008), it clearly concludes that there is substantial potential for interline connectivity to be exploited by passengers. The previous contributions, however, are exclusively based on airlines schedules data and hence, there is room to improve on existing methods by adding information on actual itineraries flown by passengers. An exploratory study on this topic will be particularly valuable in the current early stages of implementation by airports and airlines worldwide.

In this context, we aim to evaluate the potential for widespread development of this type of traffic and the implications for airports and airlines. To that end, we use a Quality of Service Index (QSI) methodology, based on coefficients obtained using a zero-inflated Poisson regression on global MIDT data for June 2014, to forecast the amount of traffic that could be captured by self-connecting travel alternatives created using a connections builder (CB) algorithm. We identify the airports that have the highest potential to facilitate self-connections and investigate (with a second-stage regression) the role that geographic location, destination mix, or airline dominance play in generating that potential. In addition, we discuss the factors that hinder or facilitate the necessary airline agreements at the major locations.

The rest of this paper is structured as follows: Section 2 provides a review of the literature in all subjects relevant to the paper and discusses our main contributions. Section 3 introduces the case study, describes the MIDT dataset, and the methodology, including the connections builder and Poisson regression. Section 4 presents the results and discusses their main business implications. Finally, Section 5 summarizes our findings, addresses the limitations of our model, and proposes new paths for future research.

2. LITERATURE REVIEW

We build on existing literature to justify the scope of our contributions and define our methodological process. The phenomenon of self-connectivity was firstly defined by Burghouwt (2007) as "self-help hubbing". Later, it was analysed in more detail by Malighetti et al. (2008) and hinted by Jimenez et al. (2013). More recently, Fichert and Klophaus (2016) and Maertens et al. (2016) scoped the self-connectivity market out. Using data on airline schedules, Malighetti et al. (2008) developed a method to find the quickest-travel-time itineraries[3] within the European air transport network. Their main conclusion is that the majority (two thirds) of the fastest indirect itineraries were provided outside the scope of the airline alliances. In regards to the airports with the highest potential, their results reveal that airports that are relatively "central" to the European network, in both a geographical and a topological sense, are found to be the ones with the highest potential for self-connections (e.g. Munich, Paris-CDG, Stockholm, Helsinki). Fichert and Klophaus (2016) explore the self-connecting, hubbing and codesharing options of seven European airlines and suggest that airports support of self-connections is a soft approach for connection building. Using the same type of supply data as Malighetti et al. (2008) and an SQL method, Maertens et al. (2016) also assess the potential of transfers between European LCC services. They identify Barcelona,

[3] In this paper, we refer to a travel alternative, travel option or travel itinerary as a sequence of flights between the passenger's point of origin and ultimate destination. Most origin and destination markets can be served by multiple travel alternatives, which can be either direct (non-stop) or indirect (involving at least one flight connection).

London Gatwick and London Stansted as the airports with the largest LCC transfer potential. Overall, the studies above support the hypothesis that there is indeed a large opportunity for widespread development of self-connecting travel services.

From a methodological perspective, however, we aim to advance their methods by complementing the airline schedules with data on actual passenger demand (MIDT) that provides indication of the passengers' full itineraries. This is a novel contribution to the literature on self-connectivity in airline networks. It is a relevant improvement since it allows us to assign a value to each travel option based on observed demand: 1) how many passengers do actually want to travel between any two places? 2) how strongly competitive are all available travel options that passengers have actually taken within the same market? Furthermore, we employ a second-stage regression to identify the drivers of airport self-connectivity, particularly those related to airport centrality (as suggested by Malighetti et al., 2008's results) and airline/destination mix as suggested by Gatwick Airport.

Answering to this last question is a typical challenge faced by airport marketing practitioners, particularly when working in the area of route development. In this context, Halpern and Graham (2013) note the widespread use of connection-building (CB) algorithms that feed on airline schedules and return a set of valid travel itineraries (non-stop or indirect[4]) for the selected market/s. The validity of these itineraries is ensured by introducing minimum connecting time (MCT) restrictions (Veldhuis, 1997; Burghouwt and de Wit, 2005). For simplicity, many applications employ constant MCTs depending on the type of connection (i.e., domestic or international). This is the approach implemented by Malighetti et al., (2008) when evaluating self-connectivity in the European network (a 60-minute threshold was set for intra-European connections). We aim to improve on that method by using published MCTs that are specific to the relevant airport and airlines. This will allow the CB algorithm to deliver a more precise set of valid travel itineraries. A second advantage of that approach is that the quality of airport connectivity will more closely reflect the characteristics of each airport, particularly their terminal layouts that is one of the factors behind the published MCT values[5]. A second challenge for the design of a CB algorithm is to define rules to discard unrealistic travel itineraries that should not be included within a pool of competitive travel options for passengers. To that end, the established approach is to define maximum connecting times and/or maximum geographic detours (for example, to exclude itineraries that involve backtracking). Seredyński et al. (2014) combined both types of impedances in relation to the non-stop travel time in each market (thus defining a maximum travel time increase). They showed how this approach helps the CB to return fewer unrealistic itineraries. We will implement this approach in our methodology.

A second element of our methodology is forecasting, for which there is a wide range of established methods, including gravity models (Grosche et al., 2007), logistic regression (Coldren et al., 2003; Coldren and Koppelman, 2005), Neural Networks (e.g., Zhang and Qi, 2005), or the Quality of Service Index-QSI (Tembleque-Villalta and Suau-Sanchez, 2015). According to Halpern and Graham (2013), QSI models have been adopted as an industry standard and are widely applied by airports to forecast market shares of new routes. In order to improve the applicability of our results in a practitioner context, we adopt the QSI method to our case study. QSI models assign a weighted "score" to each travel alternative based on a set of predictors of passenger choice. Market shares are then calculated as the ratio between the individual scores and the sum of the scores of all travel alternatives in each origin and

[4] Only flight itineraries up to two stops are typically included since demand for longer itineraries is marginal (Coldren and Koppelman, 2005).

[5] Airports with multiple terminals tend to have longer MCTs.

destination market. Wei and Hansen (2006) highlight the arbitrary nature of predictor weights on most QSI applications as a common criticism levied on these models. In order to address this shortcoming, we calibrate the variable weights using an econometric regression on MIDT data. In accordance with the nature of passenger bookings, we model it as count data and employ both Poisson and negative binomial regression models (Mao et al., 2015). These have been used for air travel demand forecasting (Johansson et al., 2011; Mao et al., 2015) as well as in areas like airline reservation systems (Lee, 1990) and airline safety (Madsen et al., 2015). Our case study, however, is different from previous applications in which we also need to model the absence of demand for the majority of valid itineraries available in air transport networks. We employ zero-inflated models to that end, which is a first in this subject area.

The most common predictors of air travel demand include fares, frequencies (typically over an average week), connecting times, number of stops, travel detours (both geographical and in terms of travel time), aircraft type, booking class (e.g. business vs. economy), or departure time (Tembleque-Villalta and Suau-Sanchez, 2015). All these variables will be included in our regression model, except booking class, for which information is not available. The broader literature on airport choice has also established the importance of airport surface accessibility (Ashford and Bencheman, 1987; Windle and Dresner, 2002; Pels et al., 2003; Johnson et al., 2014; Hess and Polak, 2005; Lieshout et al., 2015). However, we cannot account for this variable either due to the lack of information about the place of residence of passengers.

This paper also contributes to the literature on LCC interline agreements, which already includes theoretical (Kawamori and Lin., 2011) and empirical studies (Morandi et al., 2014). Past contributions focus mainly on alliance membership and codesharing, with no previous discussions on self-connection agreements. However, we argue that, due to the similarities with other types of interline collaboration, the incentives of LCCs to enter into these agreements must be linked to the same factors that explain the existence of LCC codeshares, which were investigated by Morandi et al. (2014). These authors highlight the positive impact of reciprocity, which is likely to be a relevant factor for self-connections as well since the feeding airline (the airline that serves the first leg) assumes the risk of passenger compensation if its flight is delayed. The lack of symmetry in traffic flows (which we will be able to measure with our data) may result in an unbalanced distribution of risk between both airlines. We will build part of our exploratory discussion on potential self-connection agreements on this concept.

In summary, we contribute to the literature on self-connectivity in airline networks by employing data on passenger demand and published minimum connecting times to improve the quality of our connectivity and forecasting models. While we use established methods (such as CB, QSI, and econometric regression), we adapt them to the characteristics of our case study by using zero-inflated count data models. Regarding the results, we provide the first regression analysis on the drivers of airport self-connectivity, the first ranking of airports based on demand-weighted potential for self-connectivity as well as the first discussion on the hurdles and facilitating factors of airline self-connecting agreements.

3. DATA AND METHODOLOGY

3.1. Case study and datasets

We focus on global air transport markets, as represented by the passenger itineraries included in our MIDT dataset for the first week of June 2014. Each record contains information on the ticketing airline, as well as the points of origin and destination, the connecting airports (up to two intermediate stops), and the number of passengers. In total, the MIDT dataset contains 1.4 million different itineraries in 462,599 origin-destination markets, involving slightly over 54 million passengers, 23% of which (12.3 million) travel indirectly (i.e. with at least an intermediate stop) to their destinations. Table 1 shows the distribution of this global demand for indirect air travel by geographical markets. The largest market for connections by far is the North American one, followed by Asia-Pacific and the European Economic Area (EEA). These three intra-regional markets account for 46.5% of global connections. In regards to intercontinental markets, the busiest ones connect the three aforementioned regions as well as North America with Latin America and Caribbean.

Table 1. Distribution of global passenger connections by geographical markets (June 2014)

(passengers travelling between)	*EEA*	*Rest of Europe (non-EEA)*	*Africa*	*Asia-Pacific*	*Latin America and Caribbean*	*Middle East*	*North America*
EEA	936,695	341,641	247,556	889,945	314,126	200,643	865,668
Rest of Europe (non-EEA)		178,775	23,292	118,056	16,253	52,745	90,279
Africa			88,813	158,819	17,218	81,815	110,605
Asia-Pacific				1,621,334	86,000	398,835	728,782
Latin America and Caribbean					544,963	19,651	835,993
Middle East						81,660	155,747
North America							3,172,362

Source: MIDT

Table 2 shows the top-ten intermediate hubs in the sample when all markets are considered. As expected, US hubs dominate the ranking due to their massive domestic connectivity. In addition, note the relatively high positions of airports like Dubai and Istanbul that have experienced explosive growth over the last decade. Only Frankfurt and Paris-CDG represent European airports in this top ten. Other European hubs, such as Heathrow or Amsterdam are not present due to the reduced (or inexistent) amount of domestic connections. Table 2 also shows the top-ten airlines according to global passenger bookings. Note the presence of three LCCs such as Southwest, Ryanair, and Easyjet, which supports the hypothesis that a large number of self-connecting travel alternatives will be found among these airlines that do not typically transfer flights with themselves or other airlines.

Table 2. Top-ten hub airports and airlines in the sample (1st week June 2014)

Airport	*Code*	*Originating Passengers*	*Connecting Passengers*	*Ticketing Airline*	*Code*	*Bookings*
Atlanta	ATL	291,752	651,432	Delta	DL	2,346,046
Dallas/Fort Worth	DFW	255,737	404,912	Southwest	WN	2,220,766
Chicago O'Hare	ORD	370,466	378,298	United Airlines	UA	2,036,182
Frankfurt	FRA	261,662	368,446	Ryanair	FR	1,833,086
Dubai	DXB	266,684	342,911	American Airlines	AA	1,488,161
Charlotte	CLT	103,261	322,001	Easyjet	U2	1,274,616
Istanbul Ataturk	IST	281,084	317,483	China Southern	CZ	1,182,269
Amsterdam	AMS	279,568	252,845	US Airways	US	1,128,890
London Heathrow	LHR	486,672	242,263	China Eastern	MU	1,019,959
Houston Intercontinental	IAH	172,038	241,843	All Nippon Airways	NH	848,634

Source: MIDT,

Table 3 provides summary statistics on the available price data. While the average fare of non-stop trips is lower than indirect trips, the price distributions present substantial overlap between itinerary types, thus indicating that it is possible to find indirect itineraries that undercut their non-stop counterparts. The quality of the price data is not optimal tough. Only 28.8% of itineraries in the MIDT dataset include price information. This poor coverage limits the role that airfares can play in our QSI model.

Table 3. Distribution of prices per type of itinerary (1st week June 2014)

USD	*Mean*	*s.d.*	*Min*	*Max*
non-stop fares	271.27	261.60	10.00	7,312.00
1-stop fares	464.25	568.52	10.00	19,803.00
2-stop fares	605.26	697.84	10.00	18,918.00

Source: MIDT

The original sources of information for the MIDT dataset are Global Distributions Systems (GDSs) such as Galileo, Sabre, or Amadeus, among others. MIDT is a common data source in the construction of airport route development reports (Halpern and Graham, 2013). The raw data on the passenger bookings has been adjusted by the provider (OAG Traffic Analyser) using a proprietary algorithm based on frequencies and supplied seats in each flight sector. The reliability of these adjustments has been discussed in past studies that also employed these types of datasets for connectivity purposes (e.g, Suau-Sanchez et al., 2015, 2016).

The CB algorithm is built on a dataset of global flight schedules during the first week of June 2014, whose primary source is the OAG Schedules dataset. After simple data processing, the supply dataset comprises 655,987 unique records of scheduled passenger flight departures for 747 airlines that offered 89.7 million seats across a network of 2,998 commercial airports. Each record indicates the operating airline, alliance membership (if applicable), flight number, origin and destination airport codes, aircraft type, number of seats, flight distance, departure and arrival times, and departure and arrival terminals (if applicable). Finally, we also have a dataset of minimum connecting times, including the airports' default values plus around 68,000 airline-specific exceptions. This was obtained from the OAG Connections Analyser.

3.2 Connection Builder

The passengers' choice of itinerary is affected by many variables, one of which is flight frequency. This is measured using a CB method that finds all valid travel itineraries for the observed origin and destination markets. For each MIDT record, a search is made in the schedules dataset for all valid flight combinations from the origin airport/city to the destination airport/city (up to a maximum of two stops). Airports in multi-airport systems are considered mutual substitutes and hence pooled together as they belong to the same market. The reason for this is that self-connecting passengers tend to be price-sensitive passengers (Burghouwt, 2007; Maliguetti et al., 2008; Fageda et al., 2015; OAG, 2016). They are usually footloose and look for the ground transport option (Hess and Polak, 2005) and airport that allows them to complete the overall itinerary in the most affordable way.

No interline restrictions are imposed through flight connections within the same alliance as flagged as such. For a flight combination to be valid, it must meet the published minimum connecting times[6].

CB should also be able to discriminate between realistic and unrealistic flight combinations (Redondi et al., 2011; Seredyński et al., 2014; Grosche and Klophaus, 2015). To that end, we impose a maximum allowable geographic detour for each market (ratio between indirect and non-stop flight distance). In order to mitigate the influence of outliers (exceptionally large detours are observed for a negligible share of passenger itineraries), we discard every flight combination found by the CB algorithm that is above the 95% percentile of the market-specific distribution of geographic detour calculated from the itineraries in the MIDT file. An additional constraint is imposed in regards to maximum travel time increase (ratio between total indirect travel time, including flight connections, and non-stop travel time[7]). That limit is established

[6] Previous studies consider a standard minimum connecting time of 45 minutes for all airport (See for example Veldhuis (1997), Burghouwt and de Wit (2005), Burghouwt (2007), Suau-Sanchez and Burghouwt (2012) or Maertens et al. (2016)). The application of a clear and consistent airport-specific criterion for minimum connecting times (i.e., the published minimum connecting times) allows us to provide a more transparent appraisal comparable across markets. We acknowledge, however, that the implementation of a self-connectivity platform could involve a modification of the minimum connecting times. For example, ViaMilano only connects flights with a MCT of 90 minutes.

[7] In the absence of direct flights, a hypothetical non-stop travel time is calculated using average travel speeds for the relevant aircraft, trip distance, and direction of travel.

at the 95% percentile in the distribution of travel time increase. This distribution includes all passenger bookings in traditional flight combinations within a one-hour window with respect to the best weekly indirect travel time in the market[8].

Table 4 shows the top ten airports in regards to the number of potential flight combinations identified by the CB algorithm under the above restrictions. Initially, a flight combination is labelled as "self-connecting" if either: 1) both arriving and departure airlines are LCCs (following the list provided by ICAO, 2014)[9], or 2) arriving and departure airlines are not part of the same alliance (including all subsidiaries of partner airlines). This broad definition, however, leaves some traditional flight connections misclassified, such as those provided in virtue of out-of-alliance interlining agreements and the transfer services already provided by LCCs at selected locations. We identify these cases if by cross-checking our CB flight combinations against the published Minimum Connecting Times, as it is common that airlines providing those special connections file an exception to the airport's default values. As seen in Table 4, there is indeed a latent potential for self-connectivity in the worldwide air transport network, to be exploited primarily by US and European hubs.

Table 4. Airports with the largest number of potential connections

Airport	*Code*	*Total flight combinations: 1- and 2-stop*	*1-stop*	*2-stop*	*Self-Connecting Potential: Flight Combinations*
Chicago O'Hare	ORD	2,347,620	1,416,252	931,368	1,329,946
Atlanta	ATL	2,133,629	1,433,229	700,400	919,061
Frankfurt	FRA	1,526,101	451,983	1,074,118	537,119
Denver	DEN	1,227,234	611,557	615,677	712,287
Dallas/Fort Worth	DFW	973,254	571,342	401,912	310,151
London Heathrow	LHR	941,825	364,539	577,286	498,307
Amsterdam	AMS	910,721	288,204	622,517	313,617
Los Angeles	LAX	896,524	482,658	413,866	727,895
Houston Intercontinental	IAH	824,178	426,255	397,923	206,992
Minneapolis/St. Paul	MSP	797,839	344,292	453,547	263,320

The outcome of this stage is a dataset of 16.8 million unique itineraries that the CB identified as valid travel alternatives within the selected markets. These records include all fields that identify the individual flights taken by the passengers, such as operating airlines, flight numbers, departure and arrival times and terminals, and seat capacity.

3.3 Poisson regression and QSI model

Once we have obtained an initial set of travel itineraries for all sample markets, we combine the CB information with the passenger bookings in our MIDT dataset in order to develop a forecasting model[10]. With that model, the goal is to produce 1) a baseline scenario that represents the current situation (as of June 2014) in terms of self-connectivity, and 2) two different development scenarios that show predicted self-connecting traffic flows in the event of a widespread implementation of self-connectivity platforms, such as GatwickConnects or ViaMilano.

[8] This is meant to allow for variability across frequencies in the same itinerary (a similar arbitrary threshold was imposed by Redondi et al., 2011). To determine the 95% percentile in the distribution of travel time increase, we carry out an auxiliary computation: MIDT passenger bookings are distributed across the weekly CB frequencies - that match the MIDT itineraries- according to available seat capacity. Well-connected frequencies fill up quickly with passengers from different markets so several rounds of processing are used to re-allocate overcapacity passengers in less attractive frequencies.

[9] This applies to either flight connection in the case of 2-stop itineraries.

[10] The MIDT data only indicates the ticketing airline, which may not operate all flights in indirect passenger itineraries. Thus, we cannot simply measure self-connectivity from the MIDT file and we need to model it.

The combination of the CB flight frequencies with the MIDT weekly passenger bookings returns a dataset of 3.03 million consolidated itineraries[11]. 74.92% of the consolidated records (2,271,082) did not have any passenger bookings. These itineraries are kept in the dataset because, for our research purposes, it is crucial to understand the factors that make valid itineraries (that meet all published minimum connecting times) not capture any demand. The self-connecting nature of many of these itineraries is expected to be one of these factors.

The dependent variable is the number of weekly passenger bookings per itinerary. Table 5 provides some descriptive statistics on this variable. Note that bookings only take non-negative integer values; hence, it can be defined as count data (Mao et al., 2015). Poisson regressions are typically used to model count data. However, these models are restrictive in the sense that the Poisson distribution assumes that the conditional mean is equal to the conditional variance. This assumption is not met by our data (Table 5), which shows clear signs of overdispersion.

Table 5. Descriptive statistics of dependent variable

variable	*n*	*mean*	*variance*	*Zero obs*	*p1*	*p75*	*p90*	*p95*	*p99*	*max*
bookings	3,031,314	16.86	35,216.66	2,271,082	0	1	9	25	367	31,422

One way to deal with the high variance is to account for the distortion associated to the excessive amount of zero-booking travel itineraries. To that end, we employ a zero-inflated Poisson regression in order to separate between "true zeros" and "excess zeros" (Greene, 1994). This method models two separate data generation processes for each observation (one that generates zero counts and another generating Poisson counts), with the result of a Bernoulli trial indicating which process is used. Thus, for travel itinerary *i* (Y_i), the zero-generating process is chosen with probability φ_i and the Poisson process with probability (1- φ_i):

(1) $\mu_i = \exp(x_i'\beta)$

(2) $P(Y_i = 0|x_i, z_i) = \varphi_i(z_i'\gamma) + \big(1 - \varphi_i(z_i'\gamma)\big)\exp(-\mu_i)$

(3) $P(Y_i = y_i|x_i, z_i) = (1 - \varphi_i(z_i'\gamma))\frac{\mu_i{}^{y_i}\exp(-\mu_i)}{y_i!}$

The φ_i probability is modelled against the characteristics of each observation (z_i) using a logistic function with parameters γ to be estimated. The Poisson process has mean μ_i that is regressed against the characteristics of each observation (x_i) using a log-linear specification with parameters β to be estimated. The mean and variance of the zero-inflated Poisson model are given by:

(4) $E(y_i|x_i, z_i) = \mu_i(1 - \varphi_i)$

(5) $V(y_i|x_i, z_i) = \mu_i(1 - \varphi_i)(1 + \mu_i\varphi_i)$

As seen in Eq. 5, this model allows for overdispersion (variance is higher than the mean). An extra component of overdispersion can be introduced by employing a zero-inflated negative binomial regression, with identical formulation as above except for an extra parameter of unobserved heterogeneity for each observation, which follows a Gamma distribution with mean 1 and variance α. The conditional variance of this model is given by[12]:

(6) $V(y_i|x_i, z_i) = \mu_i(1 - \varphi_i)(1 + \mu_i(\varphi_i + \alpha))$

[11] The reason this dataset is smaller than the CB one is that CB frequencies need to be aggregated by airport codes to match the MIDT format. Since CB and MIDT records are, at this stage, aggregated at an airport code level -not at a flight number level- seat capacity restrictions are assumed to be met. Otherwise, there would be tickets in the MIDT file that could not have been sold.

[12] When α=0 the negative binomial converges into the Poisson model.

In order to estimate both models, there is need to define the *x* and *z* variables. In accordance with the previous literature, the following predictors of passenger choice are included:

1) Total weekly frequencies per itinerary: count data can be treated as the product of an incidence rate (in our case, bookings per individual frequency within each itinerary) and an exposure (frequencies per itinerary). In order to be consistent with that conceptualization, we define weekly frequencies as exposure variable, with its coefficient is restricted to 1. The interpretation of the remaining coefficients is thus made in terms of incidence rate.

2) Seat capacity at market and itinerary levels: we separate between direct and indirect (1-stop and 2-stops) seats. The reason is that the attribution of indirect capacity to a specific origin-destination market is diluted by the fact that this capacity is likely dominated by the respective non-stop markets. An indicator of concentration in seat capacity across airlines (the HHI: Hirschmann-Herfindhal Index) and the share of the ticketing airline's seat capacity to total market capacity are included as well in order to control for the effects of market dominance.

3) Number of stops and average airfares: Borrowing from Coldren and Koppelman (2005), we create a set of dummy variables that indicate whether the itinerary is non-stop, one-stop- or two-stops in comparison with best available itinerary in each market. For example, we find one-stop itineraries that operate in markets where non-stop connections are either available or not. This two-level categorisation aims to capture the diversity in competitive environments across all sample itineraries. In regards to prices, due to data restrictions traditional itineraries are given average fares per type of connection between airport-pair markets (i.e. prices are not airline-specific). Self-connecting travel options with missing fare data are given a sum of the average non-stop prices for each travel segment (as if the flight segments were bought separately). As a consequence, these price variables only aim to capture additional demand effects linked to number of stops and self-connectivity. To that end, we identify the indirect itineraries that present the best average fares in each market and also calculate the difference between an itinerary's fare and the best in the market.

4) Travel time increase (TTI): It is expected that itineraries with longer travel times (related to either geographic detour of flight transfers) are less attractive to passengers. The impact of TTI is differentiated according to number of stops.

5) Connectivity: The model accounts for two aspects of airline connectivity that can have an impact in demand. First, the proportion of self-connecting frequencies in the consolidated itinerary[13]. Second, inter-terminal connectivity labels those itineraries where a transfer between different terminals is required at any time during the trip.

6) Other: The Poisson model is completed with other common predictors of passenger choice, such as aircraft type (calculated as seats per frequency), market length (great circle distance from origin to destination), and departure time (morning: 6am-12pm; afternoon: 12pm-6pm; and evening: 6pm-12am, all times UTC). In addition, we include fixed effects for the largest hubs and airlines in order to capture any unobserved heterogeneity.

7) Zero-generating process: this equation aims to separate "excess zeros" from the "true zeros" in the sample. The way we adapt this conceptualization to our case study is the following. Excess zeros represent itineraries that were not easily accessible to passengers because of not appearing alongside traditional flight connections in reservation systems and required an extra search effort by the passenger. Thus, one can expect self-connecting itineraries to be disproportionately empty of bookings as they may be actually unknown to the air travellers that do not wish to make that effort when planning their journey. This contrasts to the role of self-

[13] Given an airport itinerary and ticketing airline, both traditional and self-connecting travel options can be found.

connectivity in the Poisson model, where it is interpreted as an impedance to informed passengers that may or may not self-connect depending on the other aspects of the itinerary. TTI is the second variable that can explain a disproportionate amount of zero-bookings for travel itineraries. This is linked to the fact that our CB model may have actually left in some unrealistic itineraries (despite all filters, this is always to be expected as argued by Seredyński et al., 2014). Including TTI in the zero-generating process helps to control for that.

The estimation output is presented in Table 6. Both equations are globally significant and the signs of the coefficients are similar. Market seat capacity is generally shown to have a negative impact on the average number of bookings per frequency within each itinerary. This is interpreted as a sign of passengers spreading out among a higher number of competitive travel options, as also indicated by the negative impact of the itinerary's capacity share. The dummy variables related to number of stops yield the expected results, 2-stop itineraries tend to have less passengers per frequency than 1-stop itineraries and the negative impact of indirect travel is exacerbated by the availability of better frequencies. Interestingly, having the lowest fares tends to boost demand only when indirect travel undercuts direct travel. As expected, price differentials have negative impact on demand in most cases. Travel Time Increases associated to indirect travel have, overall, a negative impact on the number of passengers per itinerary. Inter-terminal connections are seen as a burden by passengers, and the same applies to self-connections in 1-stop itineraries. The coefficients for departure times indicate that morning frequencies are preferred to other times. Aircraft size and great circle distance have reasonable values: larger and more comfortable aircraft boosts demand (Tembleque-Villata and Suau-Sanchez, 2015) and market length reduces the intensity of air travel demand, which is mostly intra-regional in our dataset (Table 1). As expected, self-connectivity increases the probability of an itinerary to capture zero bookings[14], and the same applies to Travel Time Increase for 2-stop itineraries. Vuong's test (Vuong, 1989) confirms the significance of the zero-inflated models with respect to the base Poisson and Negative Binomial specifications.

Table 6. Estimation output

Dependent variable: bookings	***zero-inflated Poisson***			***zero-inflated negative binomial***		
	coeff.	***s.d.***	***prob.***	***coeff.***	***s.d.***	***prob.***
Non-stop weekly seat capacity (market)	-1.14E-07	1.78E-08	0.000	-1.38E-05	3.43E-07	0.000
1-stop weekly seat capacity (market)	1.28E-07	1.12E-09	0.000	-1.29E-06	1.92E-08	0.000
2-stops weekly seat capacity (market)	-2.15E-07	4.38E-09	0.000	-1.20E-06	2.86E-08	0.000
HHI of weekly seat capacity (market)	-2.02E-02	1.28E-03	0.000	0.992240	1.09E-02	0.000
Share of weekly seat capacity to total market capacity (itinerary)	-2.92E-01	1.07E-03	0.000	-1.057182	8.83E-03	0.000
Non-stop weekly seat capacity (itinerary)	-1.42E-06	4.76E-08	0.000	5.00E-05	2.29E-06	0.000
1-stop weekly seat capacity (itinerary)	-1.79E-05	4.92E-08	0.000	-1.07E-05	2.50E-07	0.000
2-stops weekly seat capacity (itinerary)	-4.51E-04	2.52E-06	0.000	-2.61E-04	3.40E-06	0.000
1-stop itinerary in non-stop market	-2.106131	3.08E-03	0.000	-2.397251	1.28E-02	0.000
1-stop itinerary in non-stop market: Lowest fare	0.032840	9.80E-03	0.001	0.056405	2.75E-02	0.040
1-stop itinerary in non-stop market: Difference to lowest fare	-0.008993	4.51E-03	0.046	-0.028910	9.36E-03	0.002
1-stop itinerary in 1-stop market	-1.874847	2.89E-03	0.000	-2.196788	1.25E-02	0.000
1-stop itinerary in 1-stop market: Lowest fare	-0.055605	8.39E-03	0.000	0.039869	2.28E-02	0.080
1-stop itinerary in 1-stop market: Difference to lowest fare	-0.090083	4.81E-03	0.000	-0.070715	1.06E-02	0.000
2-stops itinerary in non-stop market	-2.966003	4.03E-02	0.000	-4.219023	5.95E-02	0.000
2-stops itinerary in non-stop market: Lowest fare	0.269284	1.66E-01	0.105	0.267041	1.94E-01	0.170
2-stops itinerary in non-stop market: Difference to lowest fare	-0.060706	5.56E-02	0.275	-0.001420	6.21E-02	0.982
2-stops itinerary in 1-stop market	-2.601100	3.65E-02	0.000	-3.531690	5.66E-02	0.000
2-stops itinerary in 1-stop market: Lowest fare	-0.597798	8.63E-02	0.000	-0.686731	1.25E-01	0.000
2-stops itinerary in 1-stop market: Difference to lowest fare	-0.222151	3.56E-02	0.000	-0.269850	4.69E-02	0.000
2-stops itinerary in 2-stops market	-2.099883	3.60E-02	0.000	-2.506499	5.61E-02	0.000
2-stops itinerary in 2-stops market: Lowest fare	0.334446	7.53E-02	0.000	0.321725	1.13E-01	0.004
2-stops itinerary in 2-stops market: Difference to lowest fare	0.026283	3.66E-02	0.473	-0.085622	5.04E-02	0.089
Travel Time Increase (Itinerary)	-0.349767	7.02E-04	0.000	-0.175307	2.25E-03	0.000
1-stop itinerary: Lowest TTI	-0.723232	2.97E-03	0.000	-0.588159	1.08E-02	0.000

[14] In the Poisson model, for 1-stop itineraries self-connections are approximately 21 times more probable to capture zero bookings ($e^{3.05}$). For 2-stop itineraries, the effect is 7.3 ($e^{1.98}$).

1-stop itinerary: Difference to lowest TTI	-0.747854	2.43E-03	0.000	-0.671225	8.56E-03	0.000
2-stops itinerary: Lowest TTI	-0.533219	3.66E-02	0.000	-0.527000	5.75E-02	0.000
2-stops itinerary: Difference to lowest TTI	-0.783746	3.29E-02	0.000	-1.039928	5.08E-02	0.000
Inter-terminal connection	-0.136117	9.10E-04	0.000	-0.189768	3.70E-03	0.000
1-stop itinerary: Self-Connection	-0.975686	1.22E-03	0.000	-0.748737	4.59E-03	0.000
2-stops itinerary: Self-Connection	-0.123097	5.58E-03	0.000	0.336646	1.21E-02	0.000
Morning Departure	-0.004886	5.94E-04	0.000	0.014162	4.47E-03	0.002
Afternoon Departure	-0.050817	6.25E-04	0.000	-0.017077	4.56E-03	0.000
Evening Departure	-0.118919	6.68E-04	0.000	-0.040425	4.78E-03	0.000
Great circle distance (market)	-0.000058	1.01E-07	0.000	-0.000040	5.68E-07	0.000
Average aircraft size (itinerary)	0.006813	2.93E-06	0.000	0.007270	3.45E-05	0.000
Constant	3.880323	9.68E-04	0.000	3.471244	9.43E-03	0.000
ln(total weekly frequencies per itinerary)	1	(exposure)		1	(exposure)	
+ airline effects						
+ hub effects						
Excess zeros						
1-stop itinerary: Self-Connection	3.051099	0.005856	0.000	4.927875	1.59E-02	0.000
2-stops itinerary: Self-Connection	1.987548	0.008755	0.000	2.694027	1.46E-02	0.000
1-stop itinerary: TTI	0.057389	0.002875	0.000	-0.122985	4.86E-03	0.000
2-stops itinerary: TTI	2.234952	0.005599	0.000	2.552701	9.99E-03	0.000
Constant	-2.264076	0.006761	0.000	-3.980780	1.55E-02	0.000
Overdispersion						
alpha				1.276802	2.73E-03	0.000
Observations: 3,031,314		Chi2 (156)	2.28E+08		Chi2 (156)	1.24E+06
non-zero: 760,232		Prob	0.000		Prob	0.000

The significance of the extra dispersion coefficient (alpha) is also evident from the estimation output. In spite of that, the choice of one model for the purposes of forecasting demand will be based on goodness-of-fit. To that end, we evaluate the deviation between the system-wide amount of indirect travel predicted by the model and the actual one (12.378 million passengers). For the Poisson model, the deviation is 0.07% while the negative binomial deviates in 1.5%. Hence, all forecasts will be based on the Poisson equation.

The regression coefficients are applied to the original CB travel itineraries in order to obtain detailed predictions of self-connecting traffic flows. The predicted values from the Poisson model will be interpreted as Quality of Service "scores". Market shares for each individual CB itinerary are then calculated as the ratio between the CB itinerary scores and the sum of the scores of all CB itineraries in the same origin and destination market. The total number of MIDT bookings per market are then distributed across itineraries according to their market shares to obtain a prediction of potential traffic (thus assigning a value to the individual flight connection). This leads to our baseline scenario of self-connectivity.

The first development scenario (Development 1) is obtained using the same procedure with two key changes: 1) removing the effect of self-connection coefficients in the Poisson processes, 2) increasing the price of self-connecting itineraries in 40 USD per transfer[15]. This will lead to a forecast of the amount of self-connection traffic in the event of a widespread development of platforms like GatwickConnect or Via Milano that eliminate the risk associated to self-connectivity (e.g. risk of losing onward flight, baggage transfer) and makes it more comparable to traditional connectivity. Note, however, that whilst self-connectivity risk is eliminated, these platforms still face two significant barriers or burdens. Firstly, the lack of customer awareness of their existence. Secondly, customers usually book their flights considering the origin and the destination, the choice of the intermediate hub might depend on different factors (e.g., price, airline loyalty, past experience, etc.). In other words, booking through a self-connecting platform forces the customer to change their booking routines and start the booking process thinking about the hub, rather than the origin and the destination.

[15] This is intended to match the price for self-connectivity at Gatwick (GBP 27.50). Alternative prices we also used (from USD 20 to USD 50) without a significant impact on the results.

In this regard, the second development scenario (Development 2) not only includes all of the above, but it also removes the effect of self-connectivity from the zero-generating process. This has as a goal predicting the traffic in the event all connecting/self-connecting itinerary options are available and visible to the customer during the booking process. In other words, self-connection options would be comparable to traditional connections.

3.4 Second-stage regression

A second-stage linear regression on the baseline predictions at each airport will attempt to identify the drivers of airport self-connectivity. This model includes the following variables: 1) the number of potential self-connections (*potential*) as measured in Table 2. 2) Degree centrality (*degree*): number of destinations served by the airport. 3) Betweenness centrality (*betweenness*): proportion of worldwide origin and destination pairs for which the airport lies in a shortest-path-length[16] itinerary. 4) Share of low-cost seats (*slcc*). 5) HHI of airline/alliance seat capacity (*hhiall*): this includes unallied airlines as individual entities. 6) Route diversity: difference between the 97.5% and 2.5% percentile in the distribution of flight distance at the airport. Two other variables considered, average length of haul (*alh*) and closeness centrality (*closeness*)[17], were later discarded due to high correlation with other variables.

4. RESULTS AND DISCUSSION

Table 7 provides a summary of the results for each of the three scenarios. The baseline model estimates that approximately 4% of passenger bookings in global air transport markets involve at least one self-connection. This share increases to 7% in the development 1 scenario (widespread of self-connectivity platforms) and it doubles to 15% in the development 2 scenario (self-connections fully visible in booking platforms). This growth comes at the expense of both non-stop traffic and traditional connectivity. Overall, the share of indirect air travel increases above from 22.93% to 28.77%. Table 8 indicates that the self-connections are heavily concentrated in airports located in Europe, North America, and Asia-Pacific. From the perspective of the origin and destination markets (Table 9), intra-regional self-connectivity represents approximately two-thirds of global self-connections. This is consistent with the frequency distribution of self-connectivity according to market length (great circle distance from origin to destination). While these connections are observed in all kinds of routes, the highest frequency appears in markets around 2,000 km. This is consistent with the findings by Fageda et al. (2015), which point out the increased likelihood of self-connections in short-haul routes that have more daily frequencies. As seen in Figure 1, there is a clear divide between short/medium-haul and long-haul self-connectivity. This result seems to relax the requirement of having a diverse mix of short and long-haul destinations in order to capture self-connecting passengers and suggest that there may be opportunities for specialization by focusing on either of the two primary demand segments for these services. Some airports may want to focus on short-haul touristic markets and, for others, the airline and destination mix will create long-haul self-connection opportunities. Thus, having a well-developed long-haul destination network does not seem to be a requisite for strong self-connectivity.

Table 7. Breakdown of passenger itineraries under different scenarios

Baseline (weekly traffic)			*Development 1 (weekly traffic)*			*Development 2 (weekly traffic)*		
Itinerary	***Bookings***	***%***	***Itinerary***	***Bookings***	***%***	***Itinerary***	***Bookings***	***%***
Non-stop	41,638,173	77.07%	Non-stop	40,882,499	75.67%	Non-stop	38,484,108	71.23%
Indirect Traditional	10,190,087	18.86%	Indirect Traditional	9,199,773	17.03%	Indirect Traditional	7,389,050	13.68%
Self-Connecting	2,196,296	4.07%	Self-Connecting	3,942,284	7.30%	Self-Connecting	8,151,398	15.09%
Total	54,024,556		Total	54,024,556		Total	54,024,556	

[16] Path length is measured by number of flights.

[17] Closeness centrality is defined as the inverse of the sum of the minimum distances (in terms of number of flights) between each airport and all other airports in the network.

Table 8. Geographic distribution of self-connecting traffic according to hub location (baseline scenario)

Hub location	*% self-con pax*
EEA	21.4%
Rest of Europe (non-EEA)	3.0%
Africa	1.7%
Asia-Pacific	29.2%
Latin America and Caribbean	7.7%
Middle East	3.7%
North America	33.5%

Table 9. Geographic distribution of self-connecting traffic per origin and destination market (baseline scenario)

(passengers travelling between)	*EEA*	*Rest of Europe (non-EEA)*	*Africa*	*Asia-Pacific*	*Latin America and Caribbean*	*Middle East*	*North America*
EEA	9.32%	2.37%	1.58%	4.71%	2.15%	1.51%	4.71%
Rest of Europe (non-EEA)		1.07%	0.17%	0.88%	0.13%	0.44%	0.52%
Africa			0.66%	0.76%	0.10%	0.56%	0.55%
Asia-Pacific				22.90%	0.48%	3.15%	5.37%
Latin America and Caribbean					5.62%	0.17%	4.57%
Middle East						0.62%	1.16%
North America							23.78%

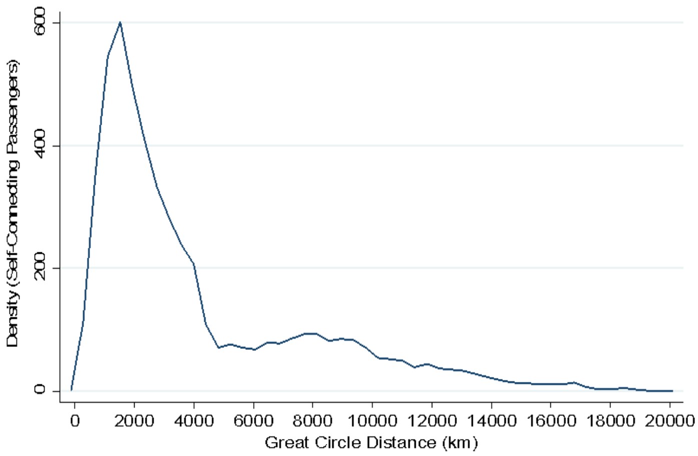

Figure 1. Distribution of self-connecting traffic according to market length (baseline scenario)

Table 10 indicates the top airports in each region according to predicted numbers of weekly self-connections under the three scenarios. The table also reports the actual number of passenger connections handled by each airport (as indicated in the MIDT file) alongside the baseline predicted values. While the model is highly accurate in replicating the system-wide amounts of direct and indirect travel, the comparison between the airport-specific values allows us to assess the accuracy of the model in a much stronger way. The average deviation between the airport-specific actual and baseline connecting traffic is 8.3%. This level of accuracy ensures that rankings based on actual and predicted traffic flows are highly consistent (rank correlation 90%+). At first sight, results indicate that the largest levels of self-connectivity are currently experienced by the airports with also the highest levels of traditional connectivity. This includes major European hubs such as Heathrow or Frankfurt, Atlanta and Chicago in the US, as well as Hong Kong and Dubai. In the European case, however, the major airports are not the ones that would benefit the most from the introduction of self-connection services. Second-tier airports with substantial LCC presence such as Barcelona, Gatwick, or Manchester experience the highest increases in connecting traffic in the development scenarios. In fact, most airports experience increases in connecting traffic with few exceptions, such as Frankfurt. Upon inspection of the raw data, this surprising result reveals a vulnerability of the primary German hub in intra-European markets, with passengers leaking from Frankfurt to alternative hubs such as Rome and Vienna.

The launch of the GatwickConnect service in late 2015 is clearly supported by our data, as it is the recently announced strategy of Ryanair to start offering connecting services at Barcelona Airport (CAPA, 2016). In Gatwick, self-connections represent almost 80% of all connectivity currently offered. The implementation of self-connection platforms is expected to increase this proportion even further and thus become the defining element of the hub experience offered at Gatwick. Regarding the particular markets served via self-connections, our analysis reveals that Gatwick can leverage its mix of airlines and destinations to capture self-connections to/from Europe's main cities as well as in North America. This contrasts with the geographical profiles of Barcelona and Copenhagen airports, whose potential for self-connectivity is mainly concentrated in Southern and Northern Europe, respectively. This suggests that patterns of hub dominance may arise as airports leverage their geographical position to capture self-connections.

The importance of geography is again evident from the second-stage regression results. Table 11 provides both standardized and unstandardized coefficients. Aside from the number of available flight connections, betweenness centrality is found to have the largest impact on the potential for self-connectivity. Airports that lie in a high proportion of shortest-path itineraries between other airports will generate more opportunities for passengers to build attractive travel itineraries on their own outside the boundaries of traditional flight connectivity. This result points at the largest hubs in each region (e.g. Frankfurt) as well as airports that serve as gateways between continents (e.g. Istanbul). As expected, the total market share of low-cost carriers operating at an airport facilitates the existence of self-connections as it does reduced airline dominance, both of which increase the viability of intra-LCC or inter-alliance transfers. On the other hand, the diversity of routes offered by an airport does not seem to have a significant impact on self-connectivity, which is linked to the aforementioned idea that airports can develop this type of traffic by focusing mostly on short-haul intra-regional markets. These conclusions complement the results obtained by previous contributions in this subject area (Malighetti et al., 2008) and serve to contextualize the views expressed by industry players in the early stages of implementation of self-connecting services (The Independent, 2015).

Table 10. Top airports according to predicted self-connectivity (weekly traffic)

Airport	*Code*	*Actual connecting pax (MIDT)*	*Baseline*			*Development 1*				*Development 2*						
			Total connecting pax	*self-con*	*% self-con*	*Total connecting pax*	*Δ pax (%)*	*self-con*	*% self-con*	*Total connecting pax*	*Δ pax (%)*	*self-con*	*% self-con*	*inline*	*iter*	*hhi*
Europe																
Frankfurt	FRA	368,446	395,396	43,477	11.0%	390,736	-1.2%	68,570	17.5%	397,965	1.9%	132,838	33.4%	0.1%	63.9%	0.013
London Heathrow	LHR	242,263	215,391	39,759	18.5%	224,715	4.3%	66,439	29.6%	253,074	12.6%	128,330	50.7%	0.0%	85.5%	0.018
Rome Fiumicino	FCO	116,467	117,439	33,404	28.4%	134,656	14.7%	58,517	43.5%	181,448	34.7%	119,564	65.9%	1.4%	62.1%	0.014
Paris CDG	CDG	235,461	193,559	31,892	16.5%	203,715	5.2%	56,712	27.8%	228,094	12.0%	111,809	49.0%	0.8%	93.8%	0.014
Amsterdam	AMS	252,845	244,110	31,047	12.7%	245,655	0.6%	50,450	20.5%	265,131	7.9%	102,040	38.5%	1.2%	0.0%	0.027
Istanbul Ataturk	IST	317,483	298,053	25,083	8.4%	295,043	-1.0%	45,876	15.5%	296,032	0.3%	91,825	31.0%	3.1%	36.8%	0.051
London Gatwick	LGW	19,618	25,067	19,830	79.1%	37,459	49.4%	33,043	88.2%	66,426	77.3%	63,154	95.1%	22.6%	44.4%	0.072
Madrid	MAD	139,298	121,985	19,618	16.1%	126,088	3.4%	33,150	26.3%	138,601	9.9%	62,166	44.9%	2.3%	75.9%	0.023
Munich	MUC	165,207	193,892	18,538	9.6%	195,323	0.7%	33,448	17.1%	211,365	8.2%	75,108	35.5%	0.0%	81.8%	0.037
Copenhagen	CPH	81,596	87,216	15,173	17.4%	93,381	7.1%	26,682	28.6%	114,357	22.5%	58,415	51.1%	1.9%	77.2%	0.045
Moscow Domod.	DME	46,540	40,533	14,780	36.5%	47,709	17.7%	24,015	50.3%	64,472	35.1%	44,711	69.3%	0.0%	0.0%	0.026
Brussels	BRU	48,388	49,059	10,692	21.8%	54,823	11.7%	19,491	35.6%	72,998	33.2%	43,779	60.0%	0.8%	0.0%	0.008
Vienna	VIE	76,528	90,152	10,588	11.7%	93,213	3.4%	19,676	21.1%	107,549	15.4%	44,591	41.5%	0.0%	0.0%	0.031
Barcelona	BCN	29,882	38,380	10,033	26.1%	42,679	11.2%	17,748	41.6%	58,971	38.2%	38,899	66.0%	2.8%	24.7%	0.023
Manchester	MAN	9,493	12,201	8,256	67.7%	15,828	29.7%	12,262	77.5%	26,679	68.5%	23,911	89.6%	3.4%	60.1%	0.005
North America																
Atlanta	ATL	651,432	702,573	67,118	9.6%	731,907	4.2%	148,694	20.3%	822,465	12.4%	349,222	42.5%	5.9%	83.1%	0.126
Los Angeles	LAX	186,786	140,442	66,126	47.1%	184,340	31.3%	121,453	65.9%	274,871	49.1%	231,852	84.3%	5.0%	92.2%	0.017
Chicago O'Hare	ORD	378,298	351,860	57,708	16.4%	380,025	8.0%	113,440	29.9%	454,046	19.5%	248,343	54.7%	0.2%	88.5%	0.167
Phoenix	PHX	176,720	162,777	51,001	31.3%	200,024	22.9%	105,789	52.9%	285,211	42.6%	218,870	76.7%	27.3%	19.3%	0.201
Denver	DEN	227,300	254,780	49,558	19.5%	275,494	8.1%	97,009	35.2%	334,435	21.4%	201,325	60.2%	0.3%	0.0%	0.106
Las Vegas	LAS	46,355	46,021	42,611	92.6%	83,711	81.9%	80,602	96.3%	168,783	101.6%	166,346	98.6%	34.7%	31.0%	0.126
New York JFK	JFK	133,871	89,648	38,851	43.3%	113,202	26.3%	69,004	61.0%	159,454	40.9%	128,163	80.4%	12.3%	82.5%	0.045
San Francisco	SFO	96,911	100,976	27,171	26.9%	115,530	14.4%	51,711	44.8%	155,130	34.3%	109,210	70.4%	3.3%	86.1%	0.035

Fort Lauderdale	FLL	26,740	24,758	22,629	91.4%	44,697	80.5%	42,792	95.7%	88,280	97.5%	86,821	98.3%	24.4%	67.3%	0.042
Dallas/Ft Worth	DFW	404,912	402,533	21,686	5.4%	388,822	-3.4%	44,540	11.5%	371,127	-4.6%	98,369	26.5%	1.4%	90.2%	0.075
Asia-Pacific																
Hong Kong	HKG	185,700	188,726	44,841	23.8%	210,164	11.4%	83,498	39.7%	273,006	29.9%	176,782	64.8%	0.0%	17.9%	0.008
Kuala Lumpur	KUL	132,999	118,216	44,621	37.7%	136,137	15.2%	75,990	55.8%	181,987	33.7%	140,441	77.2%	14.0%	53.6%	0.041
Delhi	DEL	88,946	74,075	40,120	54.2%	85,319	15.2%	58,847	69.0%	111,028	30.1%	93,208	83.9%	7.9%	55.9%	0.026
Mumbai	BOM	89,520	60,222	33,848	56.2%	72,856	21.0%	52,934	72.7%	105,142	44.3%	92,013	87.5%	10.8%	57.2%	0.029
Singapore	SIN	155,875	171,709	33,813	19.7%	178,251	3.8%	60,381	33.9%	203,021	13.9%	119,948	59.1%	2.7%	73.3%	0.006
Bangkok	BKK	124,824	125,294	20,806	16.6%	125,858	0.5%	34,522	27.4%	135,355	7.5%	67,083	49.6%	0.0%	0.0%	0.004
Beijing	PEK	133,034	137,512	17,799	12.9%	138,152	0.5%	32,210	23.3%	143,254	3.7%	62,797	43.8%	0.0%	78.1%	0.037
Tokyo Haneda	HND	116,381	113,444	17,668	15.6%	117,825	3.9%	34,863	29.6%	129,269	9.7%	68,587	53.1%	0.6%	80.6%	0.163
Tokyo Narita	NRT	96,722	91,472	17,614	19.3%	93,850	2.6%	28,741	30.6%	102,262	9.0%	54,296	53.1%	0.4%	66.4%	0.014
Sydney	SYD	66,395	78,514	16,438	20.9%	84,108	7.1%	28,429	33.8%	101,512	20.7%	57,234	56.4%	7.5%	66.3%	0.041
Africa-Middle East																
Dubai	DXB	342,911	333,533	40,881	12.3%	338,972	1.6%	72,419	21.4%	349,752	3.2%	136,908	39.1%	0.0%	65.8%	0.017
Abu Dhabi	AUH	151,026	119,356	18,303	15.3%	122,947	3.0%	31,729	25.8%	138,157	12.4%	63,045	45.6%	0.2%	66.4%	0.027
Johannesburg	JNB	69,293	50,677	11,244	22.2%	52,032	2.7%	16,255	31.2%	56,796	9.2%	28,976	51.0%	0.7%	51.3%	0.038
Doha	DOH	231,307	190,144	10,335	5.4%	184,988	-2.7%	20,704	11.2%	183,362	-0.9%	47,134	25.7%	0.0%	39.1%	0.045
Jeddah	JED	32,454	29,424	8,777	29.8%	34,081	15.8%	15,356	45.1%	44,813	31.5%	29,953	66.8%	3.6%	57.7%	0.027
Lat-America & Caribbean																
S.Paulo Guarulhos	GRU	133,106	110,404	50,098	45.4%	121,726	10.3%	72,990	60.0%	146,643	20.5%	114,814	78.3%	16.6%	62.1%	0.055
Mexico City	MEX	94,830	78,020	11,232	14.4%	82,189	5.3%	22,184	27.0%	93,697	14.0%	45,730	48.8%	24.7%	41.9%	0.087
S.Paulo Viracopos	VCP	29,861	11,403	11,090	97.3%	15,769	38.3%	15,500	98.3%	23,618	49.8%	23,435	99.2%	89.5%	0.0%	0.803
Brasilia	BSB	34,464	46,030	7,807	17.0%	44,035	-4.3%	13,714	31.1%	45,097	2.4%	25,077	55.6%	1.1%	0.0%	0.131
S.Paulo Congonhas	CGH	23,091	27,028	7,492	27.7%	28,444	5.2%	13,629	47.9%	35,499	24.8%	26,279	74.0%	33.8%	0.0%	0.286

Table 11. Second-stage regression: drivers of self-connectivity

Dependent variable: self-connecting potential	***Coef.***	***Robust s.d.***	***t***	***Prob***	***Beta***
Potential connections	0.031	0.003	9.880	0.000	0.627
Degree centrality	18.522	11.663	1.590	0.113	0.100
Betweenness centrality	262370.900	110448.400	2.380	0.018	0.251
Share of Low-Cost Carriers	1176.879	685.501	1.720	0.086	0.052
HHI Alliance	-2877.010	877.456	-3.280	0.001	-0.078
Route Diversity	0.037	0.105	0.350	0.725	0.010
Constant	174.971	557.429	0.310	0.754	
+199 country fixed-effects					

Tables 12 also provides several airport-specific indicators that aim to characterize the complexity in the implementation of self-connecting platforms from the airport perspective as well as the hypothetical airline negotiations at each location. The most obvious indicator of such complexity is the proportion of self-connections over total connections, which signals how much disruption to airport operations will be caused by the implementation of the self-connecting services. A high proportion of self-connectivity suggests that the airport may not be currently operating as a traditional hub and hence lack the necessary infrastructure and resorces to handle transfer passengers. For example, converting the 80% of connections at Gatwick (baseline levels) from being self-made to having them handled by airport staff may create a substantial pressure on airport baggage handling systems. On top of that, we also report the proportion of self-connecting passengers that would require an inter-terminal transfer (*iter*). The rates of inter-terminal transfer are significant for most airports and hence, they are an important factor to take into account while evaluating the feasibility and timescales of implementation. The variability across airports, however, suggests that, from a cost perspective, self-connecting charges should be different depending on the size and complexity of the airport's terminal layout. This is a factor that airports without inter-terminal transfers (e.g. Vienna) could exploit to achieve a pricing advantage.

From the airline perspective, there is a clear divide between LCC-dominated and other airports as the first translates into a higher share of inline self-connectivity. This would allow for an initial implementation of these services that is not fully dependent on interline negotiations. This, again, points at Gatwick as an ideal location to promote self-connectivity (Table 12). Similar conditions are present in Las Vegas or Ft. Lauderdale airports. The complexity of airline negotiations, in any case, will benefit from a reduction in the number of actors involved. We characterize that by calculating the Hirschmann-Herfindahl Index (HHI) on the interline traffic flows: the higher the HHI the more concentrated is interlining activity among fewer airlines. Our results show that airports like Vienna, Copenhagen, or Phoenix may benefit from a higher concentration in self-connecting frequencies, and thus simpler negotiations, in comparison with other airports of the same geographical region.

Table 12 provides information on the top airlines according to self-connecting traffic in the three scenarios. As in the airport case, the airlines are grouped in geographical regions as per their home countries. Results indicate that LCCs like Ryanair, Easyjet, Southwest, and Varig are the ones with the highest baseline levels of self-connectivity in each region as well as the highest potential for growth in the development scenarios. In spite of that, we also find traditional network carriers such as Delta or British Airways highly ranked. Since self-connectivity occurs between a feeding airline (that operates the first flight) and an onward airline, it is worth looking at the potential onward partners for each major carrier (Table 13). The first conclusion is that Southwest has the easiest implementation due to the largest proportion of inline self-connectivity (33.29%). Among the other LCCs we can mention, in descending order, Ryanair (23.78%), Easyjet (19.18%), and Jetblue (16.89%). This table also reveals the existence of a high degree of reciprocity between some airlines, for example, Delta and Southwest would be one another's top onward partner. On the European side, there is also good reciprocity between Easyjet and Air France. While there may be other considerations that prevent airlines from entering into interline agreements with direct competitors (e.g. American and United), the existence of reciprocity can be expected to facilitate negotiations. In this case, it can serve to achieve a more balanced distribution of risk between the partners (we argue that there is a higher risk for the feeding airline in regards to passenger compensation if its flight is delayed), as well as in terms of shared seat capacity. Thus, the values presented in Table 13 could be used as a guideline on potential self-connection partners.

Table 12. Top airlines according to predicted self-connectivity (weekly traffic)

Airline	*Code*	*Baseline*			*Development 1*				*Development 2*			
		Pax. Connections[18]	*Self-connect*	*% conn.*	*Pax. Connections*	*Increase conn (%)*	*Self-connect*	*% conn.*	*Pax. Connections*	*Increase conn. (%)*	*Self-connect*	*% conn.*
North America												
Southwest	WN	985,273	299,205	30.4%	1,244,895	26.4%	626,940	50.4%	1,847,922	48.4%	1,359,623	73.6%
Delta	DL	2,513,203	196,495	7.8%	2,515,693	0.1%	408,184	16.2%	2,581,624	2.6%	910,781	35.3%
United	UA	1,818,497	190,918	10.5%	1,855,262	2.0%	388,767	21.0%	1,979,067	6.7%	845,909	42.7%
American Airlines	AA	1,426,279	128,853	9.0%	1,435,344	0.6%	267,914	18.7%	1,501,922	4.6%	591,508	39.4%
US Airways	US	1,245,738	81,616	6.6%	1,229,430	-1.3%	168,995	13.7%	1,226,529	-0.2%	371,081	30.3%
JetBlue	B6	84,122	75,885	90.2%	156,502	86.0%	149,178	95.3%	312,690	99.8%	306,967	98.2%
Air Canada	AC	415,104	48,768	11.7%	419,380	1.0%	88,701	21.2%	435,690	3.9%	176,058	40.4%
Airtran Airways	FL	46,459	45,849	98.7%	105,739	127.6%	105,189	99.5%	249,065	135.5%	248,628	99.8%
Spirit Airlines	NK	44,396	43,245	97.4%	91,628	106.4%	90,651	98.9%	196,848	114.8%	196,140	99.6%
Alaska Airlines	AS	212,297	35,743	16.8%	232,192	9.4%	72,039	31.0%	281,512	21.2%	151,908	54.0%
Europe												
Easyjet	U2	75,270	75,270	100.0%	139,506	85.3%	139,506	100.0%	301,354	116.0%	301,354	100.0%
Lufthansa	LH	898,005	65,887	7.3%	880,556	-1.9%	117,175	13.3%	873,305	-0.8%	252,649	28.9%
British Airways	BA	388,800	60,624	15.6%	406,846	4.6%	112,370	27.6%	459,410	12.9%	234,125	51.0%
Air France	AF	373,577	45,958	12.3%	383,851	2.8%	85,297	22.2%	414,598	8.0%	176,982	42.7%
Ryanair	FR	42,879	42,879	100.0%	73,337	71.0%	73,337	100.0%	148,340	102.3%	148,340	100.0%
KLM	KL	353,353	37,349	10.6%	353,206	0.0%	63,931	18.1%	368,156	4.2%	133,448	36.2%
SAS	SK	302,838	33,490	11.1%	310,110	2.4%	60,112	19.4%	336,298	8.4%	125,156	37.2%
Norwegian	DY	110,007	29,191	26.5%	128,723	17.0%	53,278	41.4%	175,747	36.5%	110,772	63.0%
Alitalia	AZ	187,724	28,097	15.0%	196,368	4.6%	51,872	26.4%	223,462	13.8%	106,930	47.9%

[18] Passenger connections, in this table, are treated as "enplanements". Thus, indirect passengers are counted more than once, even if they travel within the same airline.

Vueling	VY	77,026	24,798	32.2%	95,364	23.8%	49,610	52.0%	147,762	54.9%	111,456	75.4%
Rest of World												
VARIG-gol Airlines	G3	135,313	70,354	52.0%	171,053	26.4%	117,266	68.6%	252,938	47.9%	214,036	84.6%
IndiGo Air	6E	58,173	58,173	100.0%	91,705	57.6%	91,705	100.0%	161,077	75.6%	161,077	100.0%
Azul Airlines	AD	51,673	50,900	98.5%	74,095	43.4%	73,500	99.2%	118,936	60.5%	118,603	99.7%
China Southern	CZ	469,604	46,632	9.9%	475,364	1.2%	96,990	20.4%	512,112	7.7%	215,936	42.2%
Emirates	EK	550,340	46,251	8.4%	540,663	-1.8%	81,619	15.1%	513,609	-5.0%	157,110	30.6%
Air China	CA	302,803	43,000	14.2%	323,887	7.0%	91,880	28.4%	388,189	19.9%	209,767	54.0%
AirAsia	AK	42,166	42,166	100.0%	71,874	70.5%	71,874	100.0%	133,874	86.3%	133,874	100.0%
Jet Airways India	9W	111,616	40,820	36.6%	123,930	11.0%	68,973	55.7%	163,568	32.0%	126,548	77.4%
Air India	AI	122,738	39,903	32.5%	129,989	5.9%	66,696	51.3%	160,516	23.5%	118,316	73.7%
China Eastern	MU	369,333	38,351	10.4%	378,276	2.4%	80,683	21.3%	415,615	9.9%	180,552	43.4%

Table 13. Top-10 onward airlines for the busiest feeding airlines (development 2 scenario)

Feeding	*Southwest*		*Delta*		*United*		*American Airlines*		*Jetblue*	
Onward (% self-connections)	Southwest	33.29%	Southwest	23.26%	American Airl.	26.46%	United Airl.	36.89%	JetBlue	16.89%
	Delta	18.67%	Airtran	16.42%	Southwest	18.80%	Delta	19.82%	Delta	16.20%
	US Awys.	13.11%	United	15.08%	Delta	17.47%	Southwest	10.66%	Southwest	12.04%
	United	12.36%	American Airl.	12.73%	US Awys.	6.45%	Spirit Airl.	4.29%	United	10.67%
	American Airl.	4.90%	US Awys.	10.70%	Frontier Airl.	4.52%	JetBlue	2.71%	American Airl.	9.32%
	Airtran Awys.	3.08%	JetBlue	3.10%	Alaska Airl.	2.97%	Alaska Airl.	2.13%	US Awys.	9.30%
	Alaska Airl.	2.79%	Spirit Airl.	2.19%	Virgin America	2.50%	Virgin America	2.05%	Spirit Airl.	3.42%
	Frontier Airl.	2.51%	Alaska Airl.	1.63%	JetBlue	2.48%	Air Canada	1.54%	Virgin America	1.42%
	JetBlue	1.97%	Air Canada	1.42%	Spirit Airl.	2.36%	Lufthansa	1.28%	Alaska Airl.	1.39%
	Spirit Airl.	1.56%	Frontier Airl.	1.28%	British Awys.	2.13%	Air France	1.04%	British Awys.	1.24%
Feeding	*Easyjet*		*Lufthansa*		*British Airways*		*Air France*		*Ryanair*	
Onward (% self-connections)	Easyjet	19.18%	Air Berlin	11.20%	Lufthansa	7.81%	Easyjet	9.42%	Ryanair	23.78%
	KLM	7.80%	Condor	7.52%	United	7.31%	Lufthansa	7.78%	Alitalia	6.63%
	Air France	7.65%	British Awys.	5.53%	Easyjet	4.47%	United	3.62%	Vueling	6.60%
	Alitalia	5.67%	Air France	4.75%	Delta	4.31%	British Awys.	3.48%	Easyjet	6.32%
	British Awys.	4.71%	Delta	4.26%	Air Canada	4.01%	American Airl.	3.36%	Iberia	4.01%
	Norwegian	3.38%	Alitalia	2.98%	SAS	3.89%	SAS	3.07%	Air Europa	3.99%
	Ryanair	2.95%	US Awys.	2.84%	Aer Lingus	3.28%	Turkish Airl.	2.81%	Aer Lingus	3.34%
	Lufthansa	2.81%	Emirates	2.67%	South African Awys.	3.22%	Air Canada	2.44%	Lufthansa	3.04%
	Vueling Airl.	2.70%	Vueling Airl.	2.50%	Emirates	3.11%	Vueling	2.40%	TAP Portugal	2.73%
	SAS	2.35%	Aeroflot	2.38%	Turkish Airl.	2.85%	Emirates	2.28%	Aegean Airl.	2.17%

5. SUMMARY, LIMITATIONS, AND FUTURE RESEARCH

Self-connecting passengers are those travelling with a combination of tickets where the airline/s involved do not handle the flight transfer themselves. In recent years, airports have started catering to the needs of these passengers by offering self-connection services that reduce the risks and hassle involved in such practices. This paper analyses the potential for self-connectivity in global air transport markets. In particular, we identify the airports that have the highest potential to capture self-connecting traffic, and we discuss several implementation challenges from both airport and airline perspectives. A Quality of Service Index (QSI) methodology, based on coefficients from a zero-inflated Poisson regression on MIDT data, is employed to predict the amount of potential traffic that could be captured by self-connecting travel alternatives identified with a connections builder (CB) algorithm. A second-stage linear regression examines the drivers of self-connectivity at the individual airports.

The results from the Poisson regression indicate that self-connectivity is indeed perceived as a burden, in comparison with traditional flight connections. In addition, it also increases the probability of a valid travel itinerary (which meets all minimum connecting times between the flights) to capture zero bookings. This is interpreted as a consequence of the fact that self-connecting travel options arise only out of an additional search effort undertook by informed passengers. Our baseline scenario estimates that about 4% of global air travel is currently self-connecting. This proportion is predicted to increase between three to four times if self-connectivity achieved similar characteristics to traditional flight connections with the widespread of self-connectivity platforms. Overall, self-connections are concentrated in airports and markets within Europe, North America, and Asia-Pacific. The highest frequency of self-connections is observed in markets around 2,000 km, which places the majority of self-connectivity in an intra-regional scope. Thus, having a well-developed long-haul network does not seem to be a requisite for strong self-connectivity.

Results indicate that airports that play a central role in air transport networks enjoy a larger potential for self-connectivity. Other facilitating factors are the dominance of low-cost carriers (LCCs) and airline concentration. In absolute terms, the largest amount of interline connections are still observed at primary regional hubs, though LCC-dominated airports such as Gatwick or Manchester present the highest rates of traffic growth in the event of a widespread development of self-connecting services. Results also show clear patterns of hub dominance in self-connecting routes that arise when airports leverage their geographical position, destination mix, and seat capacity to offer new indirect services in markets where there is room for additional competition.

We also investigate the conditions offered by each airport that affect the complexity of the implementation of self-connecting platforms. First, we report the proportion of self-connecting frequencies that would involve an inter-terminal transfer as a proxy for the increased pressure on the airport baggage handling systems. The rates of inter-terminal transfer are significant for most airports and hence, they are an important factor to take into account while evaluating the feasibility and timescales of implementation. From the airline perspective, there is a clear divide between LCC-dominated and other airports as the first allows for a higher share of inline self-connectivity. This would allow for an initial implementation of these services that is not dependent on interline negotiations. The complexity of those negotiations, however, will benefit from a reduction in the number of actors involved. Our results show that airports like Gatwick, Vienna, and Copenhagen may benefit from a higher concentration in self-connecting frequencies, and thus simpler negotiations, in comparison with other airports. Results also indicate that LCCs like Easyjet, Ryanair, and Southwest have the highest potential to benefit from self-connectivity, with Southwest having the simplest implementation due to the largest proportion of inline connections. However, there is also room for traditional carriers as well as an element of traffic reciprocity that is expected to increase the likelihood of successful airline partnerships.

These results, however, should be interpreted with caution and always in consideration of the limitations of our approach. Firstly, the potential agreements described in our discussion section need to be evaluated according to other factors, such as the competitive environment in the affected markets at each location. Airlines would need to consider if their self-connecting travel options may actually end up stealing passengers from direct frequencies the airline is offering elsewhere. The lack of incentives for airlines to enter in said agreement can place significant constraints on the widespread development scenarios presented in this paper. This detailed analysis is left for future research. In this regard, the implementation of such agreements and platforms is also challenged by the need of solving the awareness problems of airport-driven connection platforms (Maertens et al., 2016), the ‘baggage through-check’ problem (OAG, 2016) and increasing demands to the baggage systems that might require of additional investment. Concerning the ‘missed-connection problem’, Grimme (2008) considers that this has been overcome with the insurance included in self-connecting schemes.

Secondly, our QSI includes capacity limits at market and itinerary levels, this helps obtaining realistic forecast values, but future research could include capacity limits at the flight level, which could be more adequate for detailed airline strategy purposes.

Thirdly, the power of airports and options for airline-airport cooperation need to be reassessed. The current general view is that airports have very limited capacity to exert power and influence over airline decisions and strategies for self-connectivity. Yet, propositions like ViaMilano prove that airports can proactively find self-connectivity solutions with little airline involvement. Other airports like Chongqing Jiangbei International Airport in China provide a free hotel night stay for international passengers connecting between late-evening and early-morning flights. In this regard, cooperation between airlines and airports can help taking full advantage of this niche market, which could be mutually beneficial in those markets with limited organic growth (i.e. North America and Europe). In fact, in industry gatherings (e.g., Global Airport Development, European Aviation Conference, Passenger Terminal Expo) there is an increasing agreement on the need for closer cooperation between airlines and airports. Some airports and airlines could indeed build upon existing base of self-transfer passengers and create additional revenue streams by facilitating the connections.

Fourthly, currently there is limited knowledge on the characteristics of the self-connecting passenger. Hence, further research is also required concerning behavioural aspects. This is necessary to understand in more detail the passenger profile that self-connects or could potentially self-connect in different markets and geographical regions. This analysis could also include the issues related to customer awareness of the existence of self-connectivity platforms and how they could be overcome, as well as how established booking processes and routines could evolve and change in order to give to the hub or the intermediate airport a more prominent role in the booking of flights.

Fifthly, we did not have access to information on the actual interlining agreements between the airlines. While IATA does indeed provide that information (the Bilateral Interline Traffic Agreements – BITA document), the cost of this publication was beyond our research budget. As a result, our definition of self-connecting travel may include itineraries in which airlines from different alliances do actually provide transfer services using the airport’s default minimum connecting times. Further research should aim to implement the information on the BITA file to achieve a more precise definition of self-connecting travel.

Our analysis could also underestimate the self-connectivity potential between late-evening and early-morning flights. As mentioned in the Introduction, in some LCC airport bases is common to see price-sensitive passengers sleeping in the terminal to self-transfer between late-evening and early-morning flights. This is a phenomenon taking place in particular airports that might require of a specific analysis.

The estimation process will also benefit for higher-quality price and booking class information. This would allow for a better characterization on the impact of reduced fares (the most important factor for self-connecting passengers) on passenger demand, and to obtain an estimation on potential cost savings for passengers and revenue implications for airlines at both network and route levels. Any generation of new demand as a result of the availability of new frequencies in previously unserved markets, or seat capacity limitations at the flight level are not modelled either. With sufficient data, the QSI method can be expanded to include additional variables, such as on-time performance at a flight level. This indicator is bound to be of relevance, not only for the self-connecting passengers (risk of losing the onward flight), but also for the airlines (risk of having to pay compensation) and can bring another level of complexity to the airline negotiation process.

ACKNOWLEDGMENTS

The authors would like to thank Kevin O'Connor and the two anonymous reviewers for their comments and suggestions.

REFERENCES

Ashford, N., Bencheman, M., 1987. Passengers' choice of airport: an application of the multinomial logit model. In: Transportation Research Board, Transportation Research Record 1147: Air Transportation Issues. Transportation Research Board, Washington, DC, pp. 1–5

Burghouwt, G., 2007. Airline Network Developments in Europe and its Implications for Airport Planning. Ashgate, Aldershot.

Burghouwt, G., de Wit, J., 2005. Temporal configurations of European airline networks. Journal of Air Transport Management 11, 185–198.

CAPA, 2016. Ryanair transfer traffic & interlining; closing the gap with FSCs on product, but not on costs. Centre for Aviation. 20/04/2016.

Coldren, G.M., Koppelman, F.S., Kasturirangan, K., Mukherjee, A., 2003. Modelling aggregate air-travel itinerary shares: logit model development at a major US airline. Journal of Air Transport Management 9, 361–369.

Coldren, G.M., Koppelman, F.S., 2005. Modeling the competition among air-travel itinerary shares: {GEV} model development. Transportation Research Part A 39 (4), 345-365.

De Poret, M., O'Connell, J., Warnock-Smith, D., 2015. The economic viability of long-haul low cost operations: evidence from the transatlantic market. Journal of Air Transport Management 42, 272-281.

De Wit, J.G., Zuidberg, J., 2012. The growth limits of the low cost carrier model. Journal of Air Transport Management 21, 17-23.

Dobruszkes, F., 2013. The geography of European low-cost airline networks: a contemporary analysis. Journal of Transport Geography 28, 75–88.

Fageda, X., Suau-Sanchez, P., and Mason K., 2015. The evolving low-cost business model: Network implications of fare bundling and connecting flights in Europe. Journal of Air Transport Management 42, 289-296.

Fichert, F., Klophaus, R., 2016. Self-connecting, codesharing and hubbing among European LCCs: From point-to-point to connections? Research in Transportation Business & Management. DOI: 10.1016/j.rtbm.2016.07.001.

Gatwick Airport, 2015. http://www.gatwickairport.com/at-the-airport/flight-connections/gatwick-connects/

Graham, A., Nenem, S., Dennis, N., 2013. A method for developing a consumer preference centric airline schedule quality metric. In: Proceedings of the Air Transport Research Society Annual World Conference, Bergamo, June 2013.

Greene, W., 1994, Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models. Working paper EC-94-10, Department of Economics, Stern School of Business, New York University.

Grimme, 2008. Low cost carrier connecting flights and interlining – A conundrum finally solved? Proceedings of the 12th Annual Conference of the Air Transport Research Society, Athens, 2008.

Grosche, T., and Klophaus, R., 2015. Hubs at risk: Exposure of Europe's largest hubs to competition on Transfer city Pairs. Transport Policy 43, 55–60.

Grosche, T., Rothlauf, F., and Heinzi, A., 2007. Gravity models for airline passenger volume estimation. Journal of Air Transport Management 13, 175–183.

Halpern, N., and Graham, A., 2015. Airport route development: A survey of current practice. Tourism Management 46, 213-221.

Hess, S., and Polak, W., 2005. Mixed logit modelling of airport choice in multi-airport regions. Journal of Air Transport Management 11, 59–68.

ICAO, 2014. List of Low-Cost Carriers. http://www.icao.int/sustainability/Documents/LCC-List.pdf

Jimenez, E., Claro, J., Pinho de Sousa, J., 2013. The airport business in a competitive environment. European Journal of Transport and Infrastructure Research 13, 315-335.

Johansson, M., Arana-Vizcarrondo, N., Biggerstaff, B., Staples, J., Gallagher, N., and Marano, N., On the treatment of airline travelers in mathematical models. PLoS ONE 6 e22151.

Johnson, D., Hess, S., and Matthews, B., 2014. Understanding air travellers' trade-offs between connecting flights and surface access characteristics. Journal of Air Transport Management 34, 70-77.

Kawamori, T., Lin, M.H., 2011. Airline Alliances with Low Cost Carriers. Available at: SSRN 1798298.

Klophaus, R., Conrady, R., and Fichert, F., 2012. Low cost carriers going hybrid: evidence from Europe. Journal of Air Transport Management 23, 54-58.

Lee, A. 1990. Airline reservations forecasting: probabilistic and statistical models of the booking process. PhD Thesis. Massachussets Institute of Technology, September 1990.

Madsen, P., Dillon, R., and Tinsley., C., 2015. Airline Safety Improvement Through Experience with Near-Misses: A Cautionary Tale. Risk Analysis. In press.

Maertens, S., Pabst, H., Grimme, W., 2016. The scope for low-cost connecting services in Europe – Is self-hubbing only the beginning? Research in Transportation Business & Management. DOI: 10.1016/j.rtbm.2016.08.004.

Malighetti, P., Paleari, S., and Redondi, R., 2008. Connectivity of the European airport network: "Self-help hubbing" and business implications. Journal of Air Transport Management 14, 53–65.

Mao, L., Wu, X., Huang, Z., and Tatem, A., 2015. Modeling monthly flows of global air travel passengers: An open-access data resource. Journal of Transport Geography 48, 52–60.

Morandi, V., Malighetti, P., Paleari, S., and Redondi, R., 2015. Codesharing agreements by low-cost carriers: An explorative analysis. Journal of Air Transport Management 42, 184-191.

Morrell, P., 2005. Airlines within airlines: an analysis of US network airline responses to Low Cost Carriers. Journal of Air Transport Management 11 (5), 303-312.

OAG, 2016. Self-Connection: The Rise and Roadblocks of a Growing Travel Booking Strategy. OAG Reports.

Pels, E., Nijkamp, P., and Rietveld, P. 2003. Access to and competition between airports: a case study for the San Francisco Bay Area. Transportation Research Part A 37, 71–83.

Redondi, R., Malighetti, P., and Paleari, S., 2015. Hub competition and travel times in the worldwide airport network. Journal of Transport Geography 19, 1260–1271.

Starkie, D., 2007. The dilemma of slot concentration at network hubs. In: Czerny, A., Forsyth, P., Gillen, D., Niemeier, H.-M. (Eds.), How to Make Slot Markets Work. Ashgate Aldershot, London.

Suau-Sanchez, P., Voltes-Dorta, A., Rodríguez-Déniz, H., 2016. The role of London airports in providing connectivity for the UK: Regional dependence on foreign hubs. Journal of Transport Geography 50, 94-104.

Suau-Sanchez, P., Voltes-Dorta, A., Rodríguez-Déniz, H., 2015. Regulatory Airport classification in the US: The role of international markets. Transport Policy 37, 157-166.

The Independent, 2015. GatwickConnects: An airport trying to make connections. 18/09/2015.

Tembleque-Villata, M., and Suau-Sánchez, P., 2015. A model to analyse the profitability of long-haul network development involving non-hub airports: The case of the Barcelona–Asian market. Case Studies in Transport Policy. In Press.

Veldhuis, J., 1997. The competitive position of airline networks. Journal of Air Transport Management 3, 181–188.

ViaMilano, 2016. http://www.flyviamilano.eu/en/viamilano-transit-service/

Vuong, Q., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57, 307–333.

Wei, W., and Hansen, M., 2006. An aggregate demand model for air passenger traffic in the hub-and-spoke network. Transportation Research Part A 40, 841–851.

Windle, R., and Dresner, M. 2002. Airport choice in multiple-airport regions. Journal of Transportation engineering, 121 (4), 332–337.

Zhang, G.P., Qi, M., 2005. Neural network forecasting for seasonal and trend time series. European Journal of Operations Research 60, 501–514.